

Inteligencia artificial

Planificación

Gispert Sánchez, Francesc-Xavier

Lorente García, Ester

Martín Alaminos, David

8 de junio de 2015

Índice

1. Introducción y descripción del problema	2
2. Modelización	4
2.1. Problema básico	4
2.2. Primera extensión	5
2.3. Segunda extensión	6
2.4. Tercera extensión	7
2.5. Cuarta extensión	7
2.6. Instancias del problema	7
3. Implementación	9
4. Resultados	10
4.1. Problema básico	10
4.2. Primera extensión	10
4.3. Segunda extensión	10
4.4. Tercera extensión	10
4.5. Cuarta extensión	10
4.6. Evolución respecto al tamaño de la entrada	11
5. Conclusión	12

1. Introducción y descripción del problema

Algunos problemas de optimización resultan computacionalmente intratables con los algoritmos exactos que se conocen a causa de su complejidad algorítmica, debida a la magnitud de la estructura combinatoria subyacente. Por ello, la inteligencia artificial utiliza técnicas que podan el espacio de búsqueda y lo inspeccionan guiadas por heurísticas para poder ofrecer soluciones aproximadas de forma sencilla.

Una familia de problemas cuya resolución se ha simplificado mucho con herramientas genéricas que automatizan este proceso de búsqueda es la de los problemas de planificación. El objetivo de estos problemas es construir un plan o secuencia de acciones que permitan llegar a un cierto objetivo partiendo de un estado inicial. Se trata, pues, de problemas de síntesis, en los que hay que construir una solución basándose en el conocimiento del dominio (en este caso, en el estado del entorno, las acciones disponibles en cada momento y el objetivo al que se desea llegar).

La comunidad de inteligencia artificial ha desarrollado, a lo largo de los años, herramientas capaces de resolver problemas de planificación independientemente del dominio. Así, el usuario de estas herramientas solo tiene que modelizar su dominio y su problema mediante un lenguaje formal de representación y el sistema se encarga de la resolución automática del problema.

En este trabajo, lo ejemplificamos mediante la resolución de un problema simple de asignación que se puede modelizar como un problema de planificación para guiar adecuadamente la construcción de la asignación. El problema consiste en la distribución de T tareas de programación entre P programadores. Cada tarea tiene asociada un nivel de dificultad (1, 2 o 3, por orden creciente) y un tiempo estimado de desarrollo (un número de horas); por su lado, cada programador tiene asociada un nivel de habilidad (1, 2 o 3, por orden creciente) y una categoría según su calidad (1 o 2, por orden decreciente). Un programador puede resolver una tarea del mismo nivel de dificultad que su nivel de habilidad en el tiempo establecido o bien una tarea con un nivel de dificultad inmediatamente superior a su habilidad añadiendo dos horas adicionales. Tras la resolución de una tarea, se crea una tarea de revisión de una hora si el programador era de primera calidad o de dos horas en caso contrario, independientemente de si el programador que la lleva a cabo tiene una habilidad inferior; esta tarea de revisión tiene la misma dificultad que la inicial y tiene que ser llevada a cabo por otro programador distinto.

En esta práctica hemos considerado la versión básica del problema y las cua-

tro extensiones propuestas, de forma incremental. El problema básico consiste en la resolución directa del problema de asignación sin tener en cuenta las tareas de revisión. En la primera extensión, se añaden estas tareas de revisión al problema. La segunda extensión pretende, además, minimizar el tiempo total de resolución de las tareas. En la tercera extensión se limita el número de tareas por programador a 2 para facilitar el trabajo en paralelo. Finalmente, en la cuarta extensión, se pretende minimizar la suma del número de personas con tareas asignadas y el tiempo total de resolución.

Se desarrolla, pues, un modelo expresado en PDDL para cada uno de los problemas considerados, analizando los elementos necesarios y las diferencias entre ellos, y se resuelven utilizando la herramienta de planificación Fast Forward. El desarrollo de la práctica, así como la estructura de este informe, está influenciado por la metodología de la ingeniería del conocimiento vista en la práctica anterior, con un desarrollo incremental y basada en distintos prototipos (aunque en este caso, debido a la simplicidad del problema, las fases de identificación del problema, conceptualización y formalización prácticamente se han unido en una sola).

En el resto de este documento, explicamos los pasos del desarrollo de nuestra solución y mostramos y analizamos los resultados obtenidos.

2. Modelización

En esta sección se describen detalladamente los elementos que aparecen en los modelos desarrollados para cada una de las versiones del problema consideradas. Como se explica en la Sección 3, el modelo para cada extensión se ha construido añadiendo funcionalidades al modelo para la extensión previa. Por lo tanto, y para evitar repeticiones innecesarias, para cada extensión solo se detallan los cambios introducidos respecto a la anterior.

2.1. Problema básico

Como se ha explicado en la Sección 1, los elementos principales del problema tratado son los programadores disponibles (con las características que definen sus habilidades) y las tareas que se deben asignar a estos (también con las propiedades que caracterizan su dificultad). Por este motivo, en nuestro modelo definimos dos nuevos tipos: tarea y programador (ambos subtipos de object). Esto permite al sistema planificador ser más eficiente, pues se restringen las posibilidades de unificación y se reduce así el factor de ramificación.

Las propiedades que caracterizan las tareas y los programadores en este problema son números enteros con los que es necesario realizar operaciones aritméticas o comparaciones (por ejemplo, para determinar si un programador dispone de la habilidad suficiente para realizar una tarea). Se ha decidido, por lo tanto, crear funciones para definirlos. En particular, se utilizan las funciones (dtarea ?t - tarea) y (ttarea ?t - tarea) para obtener el nivel de dificultad y el tiempo mínimo necesario para la realización, respectivamente, de una tarea ?t y las funciones (hprog ?p - programador) y (cprog ?p - programador) para obtener el nivel de habilidad y la calidad, respectivamente, de un programador ?p.

Para esta versión básica, no se consideran las tareas de revisión inducidas por las tareas iniciales del problema. Así, pues, las tareas se encuentran en exactamente uno de dos estados posibles en el problema: no asignadas o asignadas. Para modelizar estas posibilidades, se utiliza un predicado (servida ?t - tarea) que deberá cumplirse cuando la tarea ?t esté asignada. Este predicado es necesario para saber cuándo se ha llegado a una solución (es decir, todas las tareas están asignadas) y cuándo se puede asignar una tarea a algún programador para acercarse a una solución.

Finalmente, en esta versión básica la única acción necesaria para llevar a cabo una asignación completa de las tareas a los programadores es asignar una tarea ?t

no asignada previamente a un programador ?p capaz de ejecutarla. A tal efecto, definimos una acción realiza que tiene como parámetros (?t - tarea ?p - programador) y, si se cumple que la tarea ?t no se ha asignado todavía y que el programador ?p tiene la habilidad suficiente para realizarla, la tarea se le asigna y se marca como servida.

Observamos con esta explicación que algunos de los atributos que definen las tareas y los programadores (funciones en este modelo) no se utilizan. Aún así, estos son necesarios para las siguientes versiones del problema y, manteniéndolos en este modelo, permitimos que la entrada del problema se pueda definir prácticamente del mismo modo (es decir, una entrada para una versión del problema podría servir para otra versión con modificaciones menores). Esto nos es especialmente útil para las instancias del problema generadas automáticamente.

2.2. Primera extensión

Esta primera extensión se construye a partir del problema básico añadiendo las tareas de revisión. Estas tareas de revisión, pues, provocan diversos cambios en el modelo que comentamos a continuación.

Para empezar, en esta versión no es suficiente saber que las tareas están asignadas a algún programador: hace falta, además, considerar si su revisión también se ha asignado. Además, la asignación de la tarea de revisión tiene algunas peculiaridades distintas (por ejemplo, hay que tener en cuenta que el programador que revisa una tarea no puede ser el mismo que la ha realizado).

En este modelo, no creamos un nuevo objeto de tipo tarea para representar la tarea de revisión. En vez de eso, consideramos un mayor número de estados en los que se puede encontrar una tarea con respecto al problema: una tarea puede estar no asignada para realizar, asignada para su realización pero no para su revisión o bien asignada tanto para su realización como para su revisión. Esto se modeliza añadiendo un nuevo predicado (revisada ?t - tarea) al modelo anterior. Se crea asimismo un predicado (servida_por ?t - tarea ?p - programador) que indica el programador al cual se ha asignado la tarea (sin embargo, este nuevo predicado no sustituye a (servida ?t - tarea) por motivos de eficiencia, puesto que este último tiene un factor de ramificación menor); esto se utilizará para garantizar que la tarea de revisión la lleva a cabo un programador distinto.

De este modo, la acción realiza se modifica para incluir como efecto el uso de este nuevo predicado servida_por para marcar qué programador realizará la tarea. Para esta extensión, además, es necesario crear una nueva acción para las

tareas de revisión (puesto que ampliar la acción realiza para satisfacer también estas tareas de revisión la haría excesivamente compleja): se trata de la acción revisa. Esta admite como parámetros (?t - tarea ?p - programador) y se ejecuta cuando una tarea ?t está servida pero no revisada y con un programador ?p distinto del programador que realiza la tarea y que tiene suficiente habilidad para revisarla. En este caso, el efecto de la acción es marcar la tarea ?t como revisada para indicar que la tarea de revisión asociada ya está asignada.

2.3. Segunda extensión

En esta extensión, se empieza a tener en cuenta el tiempo de realización de las tareas. Esto significa que hay que modificar las acciones adecuadamente para poder contabilizarlo y tener en cuenta las diferentes circunstancias que incrementan el tiempo de realización o de revisión (explicadas en la Sección 1).

Añadimos, pues, una nueva función llamada (ttotal) utilizada para almacenar el tiempo total de realización de tareas (incluyendo las de revisión) de la asignación parcial calculada hasta el momento. Esta función se utiliza como valor a minimizar en esta extensión. Además, su valor deberá ser actualizado con cada acción de asignación de tareas.

Por otro lado, en esta extensión no basta con saber que hay que realizar una tarea de revisión, sino que se necesita saber también cuánto tiempo requiere esta. El tiempo de revisión de una tarea depende de la calidad del programador que la haya desempeñado previamente. Por lo tanto, una posibilidad sería ampliar la acción revisa para que usase como parámetro el programador al cual se había asignado la tarea y calcular el tiempo de revisión según la calidad de este. Sin embargo, en este caso se ha optado por la introducción de dos nuevos predicados, (por_revisar_1 ?t - tarea) y (por_revisar_2 ?t - tarea) que indican que una tarea ?t comporta una tarea de revisión de una hora o de dos, respectivamente. Se cumple uno de estos dos predicados al mismo tiempo que se cumpliría un predicado servida; por lo tanto, este último pasa a ser redundante y se elimina del modelo.

En particular, el efecto de la acción realiza debe adaptarse a estos cambios. Concretamente, cuando se ejecuta realiza con una tarea ?t y un programador ?p, se incrementa ttotal con el tiempo de realización de ?t (incluyendo las dos horas adicionales por la habilidad de ?p si cabe) y se activa uno de los dos predicados por_revisar_1 o por_revisar_2 según la calidad de ?p.

La acción revisa también se adapta del mismo modo. Por un lado, cuando se unifica esta acción con una tarea ?t y un programador ?p, en la precondi-

ción se comprueba que uno de los dos predicados (`por_revisar_1 ?t`) o bien (`por_revisar_2 ?t`) se cumpla en vez de comprobar (`servida ?t`); por otro lado, en el efecto se añade un incremento de `ttotal` en una o dos horas según el caso.

2.4. Tercera extensión

El único cambio en esta extensión respecto a la anterior es que se limita el número de tareas (incluyendo las tareas de revision) que un programador puede realizar a un máximo de dos.

Para poder tenerlo en cuenta, se utiliza una función (`nprog ?p - programador`) cuyo valor es el número de tareas asignadas al programador `?p`.

En consecuencia, las dos acciones `realiza` y `revisa`, cuando se aplican con una tarea `?t` y un programador `?p`, deben comprobar como precondition que el número de tareas asignadas a `?p` sea inferior a dos e incrementar este número (codificado mediante `nprog`) como efecto.

2.5. Cuarta extensión

En esta última extensión se intenta minimizar la suma ponderada del número total de programadores que trabajarán más el tiempo total invertido en la realización del proyecto. Con ello, se pretende minimizar ambos factores a la vez.

Para implementar esta versión, pues, hay que controlar qué programadores trabajan y cuántos son. Se añaden así un predicado (`trabaja ?p - programador`) para marcar cuando un cierto programador `?p` tiene asignada al menos una tarea y una función (`ntrabajadores`) para mantener el número de programadores empleados para la asignación de tareas hasta el momento.

Las acciones `realiza` y `revisa` actualizan sus efectos acordemente y, por lo tanto, cuando se aplican con un programador `?p` que hasta el momento no trabajaba, se marca que este trabaja y se incrementa `ntrabajadores` en una unidad.

2.6. Instancias del problema

Hasta el momento se ha explicado como se modeliza el dominio de cada versión del problema. Del mismo modo, también hay que formalizar utilizando PDDL las entradas de los problemas a resolver.

Como se ha explicado, los únicos objetos del problema son las tareas y los programadores (cada uno con su propio tipo para reducir el número de posibles

unificaciones). La entrada de un problema consiste, pues, en una enumeración de las tareas y de los programadores considerados seguida de una descripción de sus características. Es decir, se inicializan las funciones que corresponden a la dificultad y el tiempo necesario para cada tarea y a la habilidad y la calidad de cada programador. Esta parte de la descripción es común para todas las versiones del problema expuestas.

Además, algunas de las extensiones requieren de la inicialización de otras funciones que se utilizan para mantener guardados valores necesarios para construir la solución de forma adecuada. En particular, a partir de la segunda extensión hay que inicializar a cero el número de horas de realización de las tareas asignadas, a partir de la tercera extensión hay que inicializar el número de tareas asignadas a cada programador a cero y en la cuarta extensión hay que inicializar el número de programadores que trabajan en las tareas asignadas a cero, puesto que en la solución inicial no hay ninguna asignación.

El estado inicial de la planificación, pues, es una asignación vacía. Esta asignación se va completando a medida que el planificador asigna tareas (tanto las tareas iniciales como las de revisión) a los programadores. Por lo tanto, una solución final se consigue cuando todas las tareas han sido asignadas a algún programador. Esto se modeliza en la versión básica del problema (en la cual no se consideran las tareas de revisión) pidiendo que todas las tareas estén servidas y en todas las extensiones posteriores pidiendo que todas las tareas estén revisadas (puesto que, por la definición de las acciones, una tarea solo puede estar revisada cuando también está servida).

Finalmente, en algunas de las extensiones se pide que el planificador intente minimizar algunos criterios. Concretamente, en la segunda y en la tercera extensiones se intenta minimizar ($ttotal$) (que representa la suma de los tiempos de realización de todas las tareas, incluyendo las de revisión), mientras que en la cuarta extensión se intenta minimizar la suma de ($ttotal$) y ($ntrabajadores$) (que representa el número de programadores usados para realizar todas las tareas).

3. Implementación

En la Sección 2, se ha descrito de forma detallada el modelo desarrollado para cada una de las variaciones del problema tratadas; estos modelos se encuentran implementados (expresados en el lenguaje PDDL) en el código adjunto a este documento. En esta sección, pues, se comenta el proceso seguido para su implementación.

Para implementar el modelo final, se ha optado por un esquema de diseño incremental basado en prototipos, siguiendo el guión propuesto en el enunciado de la práctica. Es decir, se empezó con una solución para el problema básico propuesto, que sirvió esencialmente como toma de contacto con el lenguaje PDDL y la herramienta Fast Forward, y esta se fue ampliando poco a poco y de forma ordenada para tener en cuenta cada una de las extensiones trabajadas. Por lo tanto, a lo largo del desarrollo de la práctica se han ido construyendo prototipos, todos ellos perfectamente funcionales, cada vez teniendo en cuenta más elementos del problema hasta llegar a la última extensión.

Sin embargo, la implementación del prototipo correspondiente a cada extensión partiendo de la anterior se ha llevado a cabo prácticamente en un solo paso (o sea, sin desarrollar diversos prototipos intermedios entre una extensión y la siguiente), puesto que la estructura en extensiones propuesta en el enunciado era suficientemente sencilla (es decir, los añadidos de una extensión respecto a la previa eran tan pocos que no justificaba una metodología más compleja).

Aparte de los modelos en PDDL, también se ha desarrollado un pequeño *script* en Python (se encuentra también en el código adjunto a este documento) para la generación aleatoria de instancias de entrada de nuestro problema y se ha utilizado para los experimentos descritos en la Subsección 4.6.

4. Resultados

4.1. Problema básico

4.2. Primera extensión

4.3. Segunda extensión

4.4. Tercera extensión

4.5. Cuarta extensión

El primer juego de pruebas está compuesto por tres tareas y cinco programadores, con una selección de parámetros que da lugar a diferentes soluciones según se optimice sólo el tiempo total o sólo el número de programadores que han realizado/revisado alguna tarea.

Por lo tanto, dado que esta extensión optimiza ambos criterios, podremos observar si logra obtener una solución “intermedia”, que minimice tanto el tiempo total como el número de programadores.

Si optimizáramos sólo el tiempo, obtendríamos una planificación que tarda 21 horas e involucra a 4 programadores; si, por el contrario, optimizáramos sólo el número de trabajadores, obtendríamos una planificación de 25 horas y 3 programadores. Tal como se puede observar en la salida del programa, si minimizamos los dos criterios por igual (50 % y 50 %) conseguimos una solución con una duración de 21 horas (mejor tiempo) y 3 programadores involucrados (mejor número de trabajadores).

El segundo juego de pruebas pretende tratar el caso en que la ponderación de los dos criterios no sea la misma. Consta de cinco tareas y ocho programadores, de nuevo con parámetros seleccionados cuidadosamente para que se obtengan diferentes soluciones cuando se optimizan los criterios por separado.

Optimizando los dos criterios equitativamente (50 % y 50 %) obtenemos una planificación con un tiempo total de 37 horas y 6 programadores, pero sabemos que esta solución no ha conseguido el número mínimo de programadores, que debería de ser 5 (uno para cada tarea y revisión, siempre que se cumplan los requisitos de habilidad -como es nuestro caso).

Hay que tener en cuenta que los dos criterios están expresados en unidades diferentes, pero como que no disponemos de la información para hacer una conversión de unidades ni el planificador permite optimizar expresiones lineares, aproximamos que el número de programadores debería de tener un peso del

80 %, y el tiempo total un 20 %. De esta manera, podemos obtener una planificación que tarda 37 horas y que sólo requiere a 5 programadores.

4.6. Evolución respecto al tamaño de la entrada

5. Conclusión

En este trabajo hemos resuelto un problema práctico modelizándolo como problema de planificación y utilizando un sistema de planificación automática para su resolución. A pesar de tratarse de un problema relativamente simple, este problema nos ha permitido hacernos una idea del tipo de problemas a los que se pueden aplicar las técnicas de inteligencia artificial aprendidas y, en particular, de la capacidad expresiva de lenguajes formales como PDDL y la potencia de herramientas como Fast Forward.

En esta práctica, hemos analizado el problema de asignación de tareas de un proyecto a los programadores del equipo de desarrollo y hemos modelizado los elementos que intervienen en este problema mediante objetos, predicados y acciones de PDDL. Para ello, se ha seguido un desarrollo incremental guiado por las extensiones propuestas en el enunciado de la práctica.

De este modo, hemos podido utilizar el planificador Fast Forward para solucionar instancias del problema debidamente seleccionadas y estudiar los resultados obtenidos. Así, pues, hemos podido comprobar como los sistemas de planificación automáticos diseñados de forma genérica y sin conocimiento del dominio son capaces de ofrecer buenas aproximaciones en tiempos de ejecución muy razonables, a pesar de su facilidad de uso y su potencia expresiva.